

Cognitive validity in the testing of speaking

Dr John Field, CRELLA

Background

Cognitive validity (Glaser, 1991)



- The issue is not 'authenticity'. Clearly we cannot reproduce the circumstances of a real language event in the artificial environment of a test.
- Cognitive validity requires us to find out if the **mental processes** that a test elicits from a candidate resemble the processes that he/she would employ in real-world conditions.
- At issue: *How valid is the test as a predictor of real-life performance?*
- The notion of cognitive validity has been used to investigate whether tests of scientific thinking or logical reasoning actually tap in to the processes they are supposed to measure (rather than, e.g. relying on rote learned facts). Baxter & Glaser, 1998, Thelk & Hoole, 2006

Predictive testing



- Many high-stakes language test scores are employed predictively: e.g. to show that an individual is capable of performing in a particular job, class or academic setting.
- This places a responsibility on the test designer to ensure that the test elicits behaviour similar to the behaviour that happens in a real-world context.

Cognitive validity in speaking tests



- The cognitive validity of a test can be approached from two angles.
- A. What characterises **expert (e.g. NS) speaker behaviour**?

Is speaker behaviour in a given test representative of it?

- B. How appropriate and **well calibrated** are the cognitive demands of the tasks?
 - Do they elicit behaviour like that of real-life communication?
 - Are they appropriate to the level being targeted?

Both angles can potentially inform

- **the writing of descriptors**
 - **task content**
 - **task design**

Understanding the skill

Speaking vs writing



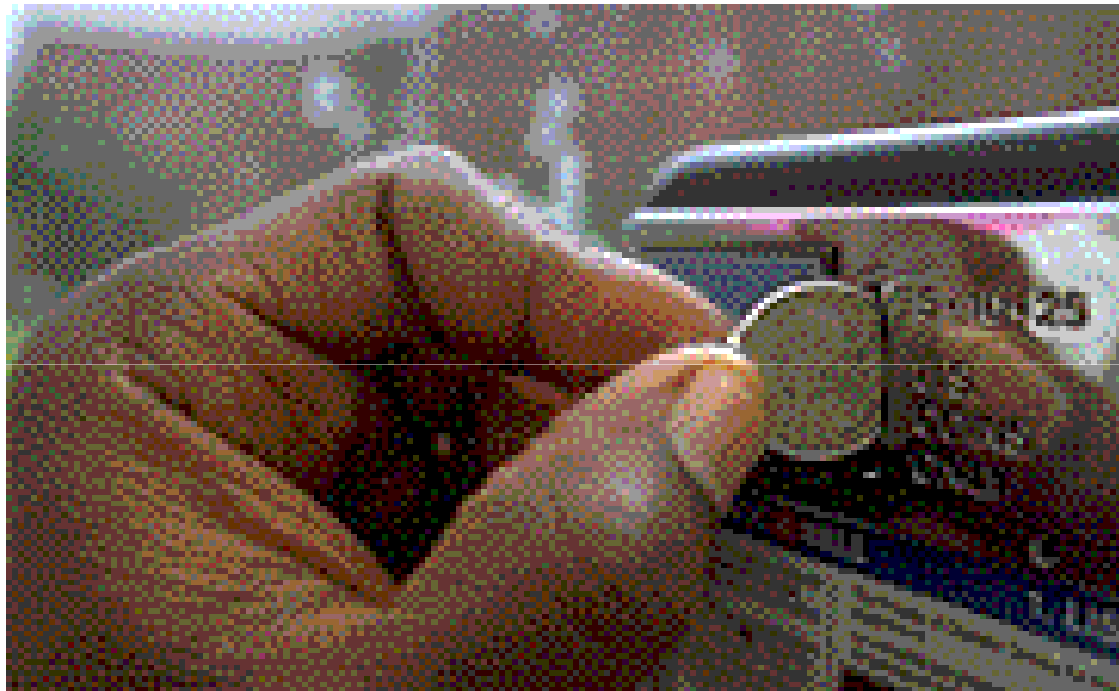
- Time pressures affecting
 - Planning
 - Execution
 - Monitoring
- Need to consider the listener reciprocally
 - Meaningful response
 - Speech rate
 - Precision of articulation
 - Accommodation

Some facts on speaking



- A speaker with a normal speech rate produces about 150 words per minute – 2.5 per second. Under pressure, the rate can rise to 5 per second.
- A normal educated adult speaker might have an active vocabulary of about 30,000 words. In fluent speech, a speaker makes the right choice from these 30,000 alternatives 2-5 times per second, and maintains the rate without any clear time limit.
- ‘There is probably no other cognitive process shared by all normal adults whose decision rate is so high.’ But the error rate is very low – about one slip per 1000 words. [based on Levelt, 1989: 199]

What's going on?

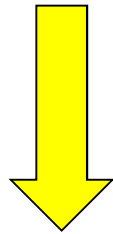


- He's putting a coin in a meter.

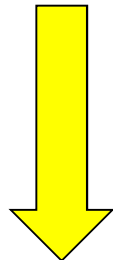
Phases of productive skills

(Levelt, 1989, 1999, Field 2011)

Abstract
meaning



Words



Sounds

Conceptualisation
Organisation

Grammatical encoding

Lexical search

Phonological encoding

Phonetic encoding

Articulation

Monitoring

Criteria for assessing a speaker



- **STANDARD L2 CRITERIA** (judgements of the listener)

Accuracy – Fluency – Complexity

- **COGNITIVE CRITERIA** (behaviour of the speaker)
- Speed of **planning** and **assembly**
- Speed and accuracy of
 - retrieval of grammar frames [PUT + object + place]
 - retrieval of lexis to fit into them [coin slot]
 - retrieval of phonological forms
 - assembly of phonetic chunks
 - Retention of speech plan in the mind
 - delivery of instructions to the articulators
- Accuracy of **self monitoring**

Conceptualisation and organisation



- Input constraints

Writing and speaking require conceptual input

- Provided by test or provided by candidate?

- Performance constraints

Most speaking events are time-constrained, with the language user obliged to process language spontaneously.

- Pre-planning may be factored in to a speaking test, but thought needs to be given to the type of real-world speaking performance that the test aims to predict.

IMPLICATIONS FOR TEST DESIGN

How do we assemble speech under time constraints?...



- Chomsky claims that language is infinitely creative. We use internalised rules to construct sentences we may never have heard before.

NP + VP + [Clause: NP+VP] + [Ind. Q clause]

I + wish + I knew + what the time is

- But speech takes place under extreme time pressures which make this process difficult to imagine.

A cognitive view of speaking expertise (Anderson, 1983)



- A novice speaker (in L1 or L2) begins with **declarative** knowledge (*knowledge that*) in the form of grammar rules or examples and lexis mediated through L1.
- Utterances demand **a high level of attention**
- This gradually turns into **procedural** knowledge (*knowledge how to*).
- a. **Composition**: Short steps become combined into larger ones
- b. **Automatisation**: Planning speech, retrieving language and delivering utterances become highly automatic and demand minimal attention

Persuasive rhetoric?



- Well I'm still proceeding with the co-operation yes because it's it's right and it's in our interests to do so and again I ++ say to people just discount those type of stories + I mean this is something we've agreed ages ago + and I think it's sensible + if for example in areas like erm + the constitution or indeed in respect of erm education it may be + or any of the issues which matter to the country + you can work with another political party because there are lots of things we have in common with the Liberal Democrats why not do it.

[Frost interview, cited Fairclough, 2000]

Persuasive rhetoric?



- Well I'm still proceeding with the co-operation yes because it's it's right and it's in our interests to do so and again I ++ say to people just discount those type of stories + I mean this is something we've agreed ages ago + and I think it's sensible + if for example in areas like erm + the constitution or indeed in respect of erm education it may be + or any of the issues which matter to the country + you can work with another political party because there are lots of things we have in common with the Liberal Democrats why not do it.

[Frost interview, cited Fairclough, 2000]

The impact of chunks



- There appears to be a breakthrough at a certain stage of an L2 speaker's development (between CEFR B1 and B2) when he/she increasingly shows signs of storing and producing chunked language.
- Yet curiously the use of formulaic chunks rarely features in the descriptors of language testing boards.
- The use of chunks enables the test taker to advance in
- **Fluency**: language is easier to retrieve
- **Accuracy**: the chunks are produced pre-assembled
- **Length of run** and **complexity**: chunks can be stitched together

[I wish I knew] + [what the time is]

An emergentist view of L2 speaker development



Progress is critically dependent upon exposure to the spoken language in terms of

- Recognising chunks
 - Phonological [shouldadone]
 - Syntactic [shoulda + past participle]
- Establishing criteria of frequency
- Analogy
 - shouldadone / mustadone / mightadone
 - shouldadone / shouldabeen / shouldathought

The impression of fluency in L1/L2: other indicators

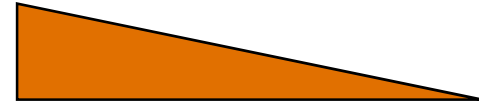


- Length of run
- Location of pauses (planning pauses vs hesitation pauses)
- Number of filled pauses
- Regularity of rhythm
- The ideal unit of planning in speech appears to be the clause.

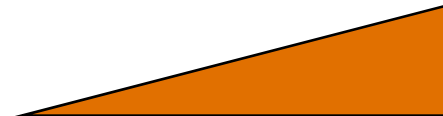
Development of strategic competence



- Amount of support



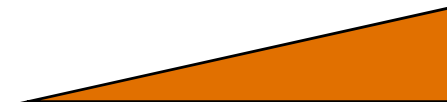
- Ability to seek repair



- Level of necessary use of compensatory strategies



- Outcomes of strategy use



Self-modification



... occurs at two points:

- While the target utterance is in the mind waiting to be articulated.

Result: hesitation

- During / immediately after an utterance

Result: false start, self-repair

- **L2 speakers (like novice L1 writers) often fail to self-monitor because they have to focus so much of their limited WM attention on planning, retrieval and articulation**

Problems of self-monitoring?

Problems of retaining a plan in the mind?



- There's a huge trust. I see it all the time when people come up to me and say 'I don't want you to let me down again'
- I think if you know what you believe it makes it a lot easier to answer questions. I can't answer your question.
- I think if you say you're going to do something and don't do it, that's trustworthiness.
- Laura and I are proud to call John and Michelle Engler our friends. I know you're proud to call him Governor. What a good man the Englers are.

Brief reflections on speaking tasks

Task demands



1. Formulation demands

- Factual presentation
- Cause /reason / example
- Compare / contrast Give opinion Negotiate
- Decision making Modality
- Summarisation
- Evaluation

Task demands



- 2 Interactional demands
 - Need to combine listening and speaking
 - Need to respond appropriately
 - Need to identify patterns of discourse
- But also some support in...
- Availability of language cues in the speech of an L1 interlocutor [what of an L2 interlocutor?]

Task demands



3. Predictability

The extent to which the task permits of lexical and syntactic **rehearsal**.

Facts about self
Limited topic area
Functional formulae

Interlocutor-led discussion
Candidate-led discussion
Candidate-led interaction

John.Field@beds.ac.uk